

19

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: Hinkkanen, Ari
- (ii) TITLE OF INVENTION: A New Fusion Protein and Its Use in an Immunoassay for the Simultaneous Detection of Autoantibodies Related to Insulin-Dependent Diabetes Mellitus
- (iii) NUMBER OF SEQUENCES: 13
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Rothwell, Figg, Ernst & Kurz, P.C.
  - (B) STREET: 555 Thirteenth Street N.W., Suite 701-E
  - (C) CITY: Washington
  - (D) STATE: D.C.
  - (E) COUNTRY: USA
  - (F) ZIP: 20004
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: US 09/015,399
  - (B) FILING DATE: 29-JAN-1998
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Ihnen, Jeffrey L.
  - (B) REGISTRATION NUMBER: 28,957
  - (C) REFERENCE/DOCKET NUMBER: 2328-111
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 202-783-6040
  - (B) TELEFAX: 202-783-6031

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: peptide
- (v) FRAGMENT TYPE: N-terminal

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Asp Tyr Lys Asp Asp Asp Asp Lys  
1 5

## (2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 amino acids

20

- (B) TYPE: amino acid
- (C) STRANDEDNESS:
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Lys Lys Lys Arg Pro Arg Lys Lys Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 10 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: C-terminal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Cys Asn Gly Ser His His His His His His  
1 5 10

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Lys Lys Lys Arg Ser Arg Lys Lys Lys  
1 5

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 979 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Arg Arg Pro Arg Arg Pro Gly Gly Leu Gly Gly Ser Gly Gly Leu  
1 5 10 15

21

Arg Leu Leu Leu Cys Leu Leu Leu Ser Ser Arg Pro Gly Gly Cys  
20 25 30

Ser Ala Val Ser Ala His Gly Cys Leu Phe Asp Arg Arg Leu Cys Ser  
35 40 45

His Leu Glu Val Cys Ile Gln Asp Gly Leu Phe Gly Gln Cys Gln Val  
50 55 60

Gly Val Gly Gln Ala Arg Pro Leu Leu Gln Val Thr Ser Pro Val Leu  
65 70 75 80

Gln Arg Leu Gln Gly Val Leu Arg Gln Leu Met Ser Gln Gly Leu Ser  
85 90 95

Trp His Asp Asp Leu Thr Gln Tyr Val Ile Ser Gln Glu Met Glu Arg  
100 105 110

Ile Pro Arg Leu Arg Pro Pro Glu Pro Arg Pro Arg Asp Arg Ser Gly  
115 120 125

Leu Ala Pro Lys Arg Pro Gly Pro Ala Gly Glu Leu Leu Leu Gln Asp  
130 135 140

Ile Pro Thr Gly Ser Ala Pro Ala Ala Gln His Arg Leu Pro Gln Pro  
145 150 155 160

Pro Val Gly Lys Gly Gly Ala Gly Ala Ser Ser Ser Leu Ser Pro Leu  
165 170 175

Gln Ala Glu Leu Leu Pro Pro Leu Leu Glu His Leu Leu Leu Pro Pro  
180 185 190

Gln Pro Pro His Pro Ser Leu Ser Tyr Glu Pro Ala Leu Leu Gln Pro  
195 200 205

Tyr Leu Phe His Gln Phe Gly Ser Arg Asp Gly Ser Arg Val Ser Glu  
210 215 220

Gly Ser Pro Gly Met Val Ser Val Gly Pro Leu Pro Lys Ala Glu Ala  
225 230 235 240

Pro Ala Leu Phe Ser Arg Thr Ala Ser Lys Gly Ile Phe Gly Asp His  
245 250 255

Pro Gly His Ser Tyr Gly Asp Leu Pro Gly Pro Ser Pro Ala Gln Leu  
260 265 270

Phe Gln Asp Ser Gly Leu Leu Tyr Leu Ala Gln Glu Leu Pro Ala Pro  
275 280 285

Ser Arg Ala Arg Val Pro Arg Leu Pro Glu Gln Gly Ser Ser Ser Arg  
290 295 300

Ala Glu Asp Ser Pro Glu Gly Tyr Glu Lys Glu Gly Leu Gly Asp Arg  
305 310 315 320

Gly Glu Lys Pro Ala Ser Pro Ala Val Gln Pro Asp Ala Ala Leu Gln  
325 330 335

Arg Leu Ala Ala Val Leu Ala Gly Tyr Gly Val Glu Leu Arg Gln Leu  
340 345 350

Thr Pro Glu Gln Leu Ser Thr Leu Leu Thr Leu Leu Gln Leu Leu Pro  
355 360 365

Lys Gly Ala Gly Arg Asn Pro Gly Gly Val Val Asn Val Gly Ala Asp  
 370 375 380  
 Ile Lys Lys Thr Met Glu Gly Pro Val Glu Gly Arg Asp Thr Ala Glu  
 385 390 395 400  
 Leu Pro Ala Arg Thr Ser Pro Met Pro Gly His Pro Thr Ala Ser Pro  
 405 410 415  
 Thr Ser Ser Glu Val Gln Gln Val Pro Ser Pro Val Ser Ser Glu Pro  
 420 425 430  
 Pro Lys Ala Ala Arg Pro Pro Val Thr Pro Val Leu Leu Glu Lys Lys  
 435 440 445  
 Ser Pro Leu Gly Gln Ser Gln Pro Thr Val Ala Gly Gln Pro Ser Ala  
 450 455 460  
 Arg Pro Ala Ala Glu Glu Tyr Gly Tyr Ile Val Thr Asp Gln Lys Pro  
 465 470 475 480  
 Leu Ser Leu Ala Ala Gly Val Lys Leu Leu Glu Ile Leu Ala Glu His  
 485 490 495  
 Val His Met Ser Ser Gly Ser Phe Ile Asn Ile Ser Val Val Gly Pro  
 500 505 510  
 Ala Leu Thr Phe Arg Ile Arg His Asn Glu Gln Asn Leu Ser Leu Ala  
 515 520 525  
 Asp Val Thr Gln Gln Ala Gly Leu Val Lys Ser Glu Leu Glu Ala Gln  
 530 535 540  
 Thr Gly Leu Gln Ile Leu Gln Thr Gly Val Gly Gln Arg Glu Glu Ala  
 545 550 555 560  
 Ala Ala Val Leu Pro Gln Thr Ala His Ser Thr Ser Pro Met Arg Ser  
 565 570 575  
 Val Leu Leu Thr Leu Val Ala Leu Ala Gly Val Ala Gly Leu Leu Val  
 580 585 590  
 Ala Leu Ala Val Ala Leu Cys Val Arg Gln His Ala Arg Gln Gln Asp  
 595 600 605  
 Lys Glu Arg Leu Ala Ala Leu Gly Pro Glu Gly Ala His Gly Asp Thr  
 610 615 620  
 Thr Phe Glu Tyr Gln Asp Leu Cys Arg Gln His Met Ala Thr Lys Ser  
 625 630 635 640  
 Leu Phe Asn Arg Ala Glu Gly Pro Pro Glu Pro Ser Arg Val Ser Ser  
 645 650 655  
 Val Ser Ser Gln Phe Ser Asp Ala Ala Gln Ala Ser Pro Ser Ser His  
 660 665 670  
 Ser Ser Thr Pro Ser Trp Cys Glu Glu Pro Ala Gln Ala Asn Met Asp  
 675 680 685  
 Ile Ser Thr Gly His Met Ile Leu Ala Tyr Met Glu Asp His Leu Arg  
 690 695 700  
 Asn Arg Asp Arg Leu Ala Lys Glu Trp Gln Ala Leu Cys Ala Tyr Gln  
 705 710 715 720

23

Ala Glu Pro Asn Thr Cys Ala Thr Ala Gln Gly Glu Gly Asn Ile Lys  
 725 730 735  
 Lys Asn Arg His Pro Asp Phe Leu Pro Tyr Asp His Ala Arg Ile Lys  
 740 745 750  
 Leu Lys Val Glu Ser Ser Pro Ser Arg Ser Asp Tyr Ile Asn Ala Ser  
 755 760 765  
 Pro Ile Ile Glu His Asp Pro Arg Met Pro Ala Tyr Ile Ala Thr Gln  
 770 775 780  
 Gly Pro Leu Ser His Thr Ile Ala Asp Phe Trp Gln Met Val Trp Glu  
 785 790 795 800  
 Ser Gly Cys Thr Val Ile Val Met Leu Thr Pro Leu Val Glu Asp Gly  
 805 810 815  
 Val Lys Gln Cys Asp Arg Tyr Trp Pro Asp Glu Gly Ala Ser Leu Tyr  
 820 825 830  
 His Val Tyr Glu Val Asn Leu Val Ser Glu His Ile Trp Cys Glu Asp  
 835 840 845  
 Phe Leu Val Arg Ser Phe Tyr Leu Lys Asn Val Gln Thr Gln Glu Thr  
 850 855 860  
 Arg Thr Leu Thr Gln Phe His Phe Leu Ser Trp Pro Ala Glu Gly Thr  
 865 870 875 880  
 Pro Ala Ser Thr Arg Pro Leu Leu Asp Phe Arg Arg Lys Val Asn Lys  
 885 890 895  
 Cys Tyr Arg Gly Arg Ser Cys Pro Ile Ile Val His Cys Ser Asp Gly  
 900 905 910  
 Ala Gly Arg Thr Gly Thr Tyr Ile Leu Ile Asp Met Val Leu Asn Arg  
 915 920 925  
 Met Ala Lys Gly Val Lys Glu Ile Asp Ile Ala Ala Thr Leu Glu His  
 930 935 940  
 Val Arg Asp Gln Arg Pro Gly Leu Val Arg Ser Lys Asp Gln Phe Glu  
 945 950 955 960  
 Phe Ala Leu Thr Ala Val Ala Glu Glu Val Asn Ala Ile Leu Lys Ala  
 965 970 975  
 Leu Pro Gln

## (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 585 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Ala Ser Pro Gly Ser Gly Phe Trp Ser Phe Gly Ser Glu Asp Gly  
 1 5 10 15

24

Ser Gly Asp Ser Glu Asn Pro Gly Thr Ala Arg Ala Trp Cys Gln Val  
20 25 30

Ala Gln Lys Phe Thr Gly Gly Ile Gly Asn Lys Leu Cys Ala Leu Leu  
35 40 45

Tyr Gly Asp Ala Glu Lys Pro Ala Glu Ser Gly Gly Ser Gln Pro Pro  
50 55 60

Arg Ala Ala Ala Arg Lys Ala Ala Cys Ala Cys Asp Gln Lys Pro Cys  
65 70 75 80

Ser Cys Ser Lys Val Asp Val Asn Tyr Ala Phe Leu His Ala Thr Asp  
85 90 95

Leu Leu Pro Ala Cys Asp Gly Glu Arg Pro Thr Leu Ala Phe Leu Gln  
100 105 110

Asp Val Met Asn Ile Leu Leu Gln Tyr Val Val Lys Ser Phe Asp Arg  
115 120 125

Ser Thr Lys Val Ile Asp Phe His Tyr Pro Asn Glu Leu Leu Gln Glu  
130 135 140

Tyr Asn Trp Glu Leu Ala Asp Gln Pro Gln Asn Leu Glu Glu Ile Leu  
145 150 155 160

Met His Cys Gln Thr Thr Leu Lys Tyr Ala Ile Lys Thr Gly His Pro  
165 170 175

Arg Tyr Phe Asn Gln Leu Ser Thr Gly Leu Asp Met Val Gly Leu Ala  
180 185 190

Ala Asp Trp Leu Thr Ser Thr Ala Asn Thr Asn Met Phe Thr Tyr Glu  
195 200 205

Ile Ala Pro Val Phe Val Leu Leu Glu Tyr Val Thr Leu Lys Lys Met  
210 215 220

Arg Glu Ile Ile Gly Trp Pro Gly Gly Ser Gly Asp Gly Ile Phe Ser  
225 230 235 240

Pro Gly Gly Ala Ile Ser Asn Met Tyr Ala Met Met Ile Ala Arg Phe  
245 250 255

Lys Met Phe Pro Glu Val Lys Glu Lys Gly Met Ala Ala Leu Pro Arg  
260 265 270

Leu Ile Ala Phe Thr Ser Glu His Ser His Phe Ser Leu Lys Lys Gly  
275 280 285

Ala Ala Ala Leu Gly Ile Gly Thr Asp Ser Val Ile Leu Ile Lys Cys  
290 295 300

Asp Glu Arg Gly Lys Met Ile Pro Ser Asp Leu Glu Arg Arg Ile Leu  
305 310 315 320

Glu Ala Lys Gln Lys Gly Phe Val Pro Phe Leu Val Ser Ala Thr Ala  
325 330 335

Gly Thr Thr Val Tyr Gly Ala Phe Asp Pro Leu Leu Ala Val Ala Asp  
340 345 350

Ile Cys Lys Lys Tyr Lys Ile Trp Met His Val Asp Ala Ala Trp Gly  
355 360 365

25

Gly Gly Leu Leu Met Ser Arg Lys His Lys Trp Lys Leu Ser Gly Val  
 370 375 380

Glu Arg Ala Asn Ser Val Thr Trp Asn Pro His Lys Met Met Gly Val  
 385 390 395 400

Pro Leu Gln Cys Ser Ala Leu Leu Val Arg Glu Glu Gly Leu Met Gln  
 405 410 415

Asn Cys Asn Gln Met His Ala Ser Tyr Leu Phe Gln Gln Asp Lys His  
 420 425 430

Tyr Asp Leu Ser Tyr Asp Thr Gly Asp Lys Ala Leu Gln Cys Gly Arg  
 435 440 445

His Val Asp Val Phe Lys Leu Trp Leu Met Trp Arg Ala Lys Gly Thr  
 450 455 460

Thr Gly Phe Glu Ala His Val Asp Lys Cys Leu Glu Leu Ala Glu Tyr  
 465 470 475 480

Leu Tyr Asn Ile Ile Lys Asn Arg Glu Gly Tyr Glu Met Val Phe Asp  
 485 490 495

Gly Lys Pro Gln His Thr Asn Val Cys Phe Trp Tyr Ile Pro Pro Ser  
 500 505 510

Leu Arg Thr Leu Glu Asp Asn Glu Glu Arg Met Ser Arg Leu Ser Lys  
 515 520 525

Val Ala Pro Val Ile Lys Ala Arg Met Met Glu Tyr Gly Thr Thr Met  
 530 535 540

Val Ser Tyr Gln Pro Leu Gly Asp Lys Val Asn Phe Phe Arg Met Val  
 545 550 555 560

Ile Ser Asn Pro Ala Ala Thr His Gln Asp Ile Asp Phe Leu Ile Glu  
 565 570 575

Glu Ile Glu Arg Leu Gly Gln Asp Leu  
 580 585

## (2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 110 amino acids
  - (B) TYPE: amino acid
  - (C) STRANDEDNESS:
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Ala Leu Trp Met Arg Leu Leu Pro Leu Leu Ala Leu Leu Ala Leu  
 1 5 10 15

Trp Gly Pro Asp Pro Ala Ala Ala Phe Val Asn Gln His Leu Cys Gly  
 20 25 30

Ser His Leu Val Glu Ala Leu Tyr Leu Val Cys Gly Glu Arg Gly Phe  
 35 40 45

Phe Tyr Thr Pro Lys Thr Arg Arg Glu Ala Glu Asp Leu Gln Val Gly  
 50 55 60

Gln Val Glu Leu Gly Gly Pro Gly Ala Gly Ser Leu Gln Pro Leu  
 65 70 75 80

Ala Leu Glu Gly Ser Leu Gln Lys Arg Gly Ile Val Glu Gln Cys Cys  
 85 90 95

Thr Ser Ile Cys Ser Leu Tyr Gln Leu Glu Asn Tyr Cys Asn  
 100 105 110

## (2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 2457 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

ACCCGCCCTC	GCCGCTCGGC	CCCGCGCGTC	CCCGCGCGTG	CCCTCCTCCC	GCCACACGGC	60
ACGCACGC	GCGCAGGGCC	AAGCCGAGGC	AGCCGCCCGC	AGCTCGCACT	CGCTGGCGAC	120
CTGCTCCAGT	CTCCAAAGCC	GATGGCATCT	CCGGGCTCTG	GCTTTGGTC	TTTCGGGTG	180
GAAGATGGCT	CTGGGGATT	CGAGAACCCC	GGCACAGCGC	GAGCCTGGTG	CCAAGTGGCT	240
CAGAAGTTCA	CGGGCGGCAT	CGGAAACAAA	CTGTGCCCC	TGCTCTACGG	AGACGCCGAG	300
AAGCCGGCGG	AGAGCGGCGG	GAGCCAACCC	CCGCGGGCCG	CCGCCCCGGAA	GGCCGCCTGC	360
GCCTGCGACC	AGAACGCCCTG	CAGCTGCTCC	AAAGTGGATG	TCAACTACCG	GTTCCTCCAT	420
GCAACAGACC	TGCTGCCGGC	GTGTGATGGA	GAAAGGCCA	CTTGGCGTT	TCTGCAAGAT	480
GTTATGAACA	TTTACTTCA	GTATGTGGTG	AAAAGTTCG	ATAGATCAAC	CAAAGTGATT	540
GATTCCATT	ATCCTAATGA	GCTTCTCAA	GAATATAATT	GGAAATTGGC	AGACCAACCA	600
CAAAATTG	AGGAAATT	TTCATGTC	CAAACAACTC	AAAAATATGC	AATTAAAACA	660
GGGCATCTA	GATACTCAA	TCAACTTTCT	ACTGGTTGG	ATATGGTTGG	ATTAGCAGCA	720
GACTGGCTGA	CATCAACAGC	AAATACTAAC	ATGTTCACCT	ATGAAATTGC	TCCAGTATT	780
GTGCTTTGG	AAATATGTCAC	ACTAAAGAAA	ATGAGAGAAA	TCATTGGCTG	GCCAGGGGGC	840
TCTGGCGATG	GGATATTTC	TCCCGGTGGC	GCCATATCTA	ACATGTATGC	CATGATGATC	900
GCACGCTTA	AGATGTTCCC	AGAAGTCAG	GAGAAAGGAA	TGGCTGCTCT	TCCCAGGCTC	960
ATTGCCCTCA	CGTCTGAACA	TAGTCATT	TCTCTCAAGA	AGGGAGCTGC	AGCCTTAGGG	1020
ATTGGAACAG	ACAGCGTGAT	TCTGATTTAA	TGTGATGAGA	GAGGGAAAAT	GATTCCATCT	1080
GATCTTGAAA	GAAGGATTCT	TGAAGCCAAA	CAGAAAGGGT	TTGTTCTTT	CCTCGTGAGT	1140
GCCACAGCTG	GAACCACCGT	GTACGGAGCA	TTTGACCCCC	TCTTAGCTGT	CGCTGACATT	1200

TGCAAAAAGT ATAAGATCTG GATGCATGTG GATGCAGCTT	GGGGTGGGGG ATTACTGATG	1260
TCCCGAAAC ACAAGTGGAA ACTGAGTGGC GTGGAGAGGG CCAACTCTGT	GACGTGGAAT	1320
CCACACAAGA TGATGGGAGT CCCTTGCGAG TGCTCTGCTC	TCCTGGTTAG AGAAGAGGGA	1380
TTGATGCAGA ATTGCAACCA AATGCATGCC TCCTACCTCT	TTCAGCAAGA TAAACATTAT	1440
GACCTGTCT ATGACACTGG AGACAAGGCC TTACAGTGC GACGCCACGT	TGATGTTTT	1500
AAACTATGGC TGATGTGGAG GGCAAAGGGG ACTACCGGGT TTGAAGCGCA	TGTTGATAAA	1560
TGTTGGAGT TGGCAGAGTA TTTATACAAC ATCATAAAAA ACCGAGAAGG ATATGAGATG		1620
GTGTTGATG GCAAGCCTCA GCACACAAAT GTCTGCTCT GGTACATTCC	TCCAAGCTTG	1680
CGTACTCTGG AAGACAATGA AGAGAGAATG AGTCGCTCT CGAAGGTGGC	TCCAGTGATT	1740
AAAGCCAGAA TGATGGAGTA TCGAACACAA ATGGTCAGCT ACCAACCCCTT	GGGAGACAAG	1800
GTCAATTCTC TCCGCATGGT CATCTCAAAC CCAGCGGCAA CTCACCAAGA	CATTGACTTC	1860
CTGATTGAAC AAATAGAACG CCTTGGACAA GATTTATAAT AACCTTGCTC	ACCAAGCTGT	1920
TCCACTTCTC TAGAGAACAT GCCCTCAGCT AAGCCCCTA CTGAGAAACT	TCCTTTGAGA	1980
ATTGTGCGAC TTCACAAAAT GCAAGGTGAA CACCACTTG TCTCTGAGAA	CAGACGTTAC	2040
CAATTATGGA GTGTCACCAG CTGCCAAAT CGTAGGTGTT GGCTCTGCTG	GTCACTGGAG	2100
TAGTTGCTAC TCTTCAGAAT ATGGACAAAG AAGGCACAGG TGTAATATA	GTAGCAGGAT	2160
GAGGAACCTC AACTGGGTA TCATTTGCAC GTGCTCTCT GTTCTCAAAT	GCTAAATGCA	2220
AACACTGTGT ATTTATTAGT TAGGTGTGCC AACTACCGT TCCCAAATTG	GTGTTCTGA	2280
ATGACATCAA CATTCCCCA ACATTACTCC ATTACTAAAG ACAGAAAAAA	ATAAAACAT	2340
AAAATATACA AACATGTGGC AACCTGTTCT TCCTACCAAA TATAAACTTG	TGTATGATCC	2400
AAGTATTTA TCTGTGTTGT CTCTCTAAAC CCAAATAAT GTGAAATGT	GGACACA	2457

## (2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 3613 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: double
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CAGCCCCCTCT GGCAGGCTCC CGCCAGCGTC GCTGGGCTC	CGGCCCCGGGA GCGAGCGCCC	60
GGAGCTCGGA AAGATGCGGC GCCCCGGCG GCCTGGGGGT	CTCGGGGGAT CGGGGGGTCT	120
CCGGCTGCTC CTCTGCCTCC TGCTGCTGAG CAGCCGCCCCG	GGGGGCTGCA GCGCCGTTAG	180
TGCCCCACGGC TGTCTATTG ACCGCAGGCT CTGCTCTCAC	CTGGAAGTCT GTATTCAGGA	240
TGGCTTGTGTT GGGCAGTGCC AGGTGGGAGT GGGGCAGGCC	CGGCCCCCTTT TGCAAGTCAC	300

CTCCCCAGTT	CTCCAACGCT	TACAAGGTGT	GCTCCGACAA	CTCATGTCCC	AAGGATTGTC	360
CTGGCACGAT	GACCTCACCC	AGTATGTGAT	CTCTCAGGAG	ATGGAGCGCA	TCCCCAGGCT	420
TCGCCCCCA	GAGCCCCGTC	CAAGGGACAG	GTCTGGCTTG	GCACCCAAGA	GACCTGGTCC	480
TGCTGGAGAG	CTGCTTTAC	AGGACATCCC	CACTGGCTCC	GCCCCCTGCTG	CCCAGCATCG	540
GCTTCCACAA	CCACCACTGG	GCAAAGGTGG	AGCTGGGCC	AGCTCCTCTC	TGTCCCCCT	600
GCAGGGCTGAG	CTGCTCCCGC	CTCTCTTGA	GCACCTGCTG	CTGCCCCCAC	AGCCTCCCCA	660
CCCTTCACTG	AGTTACGAAC	CTGCCTTGCT	GCAGCCCTAC	CTGTTCCACC	AGTTGGCTC	720
CCGTGATGGC	TCCAGGGTCT	CAGAGGGCTC	CCCAGGGATG	GTCAAGTGTG	GCCCCCTGCC	780
CAAGGGCTGAA	GCCCCCTGCC	TCTTCAGCAG	AACTGCCTCC	AAGGGCATAT	TTGGGGACCA	840
CCCTGGCCAC	TCCTACGGGG	ACCTTCCAGG	GCCTTCACCT	GCCCCAGCTT	TTCAAGACTC	900
TGGGCTGCTC	TATCTGGCCC	AGGAGTTGCC	AGCACCCAGC	AGGGCCAGGG	TGCCAAGGCT	960
GCCAGAGCAA	GGGAGCAGCA	GCCGGGCAGA	GGACTCCCCA	GAGGGCTATG	AGAAGGAAGG	1020
ACTAGGGGAT	CGTGGAGAGA	AGCCTGCTTC	CCCAGCTGTG	CAGCCAGATG	CGGCTCTGCA	1080
GAGGCTGGCC	GCTGTGCTGG	CGGGCTATGG	GGTAGAGCTG	CGTCAGCTGA	CCCCCTGAGCA	1140
GCTCTCCACA	CTCCTGACCC	TGCTGCAGCT	ACTGCCAAG	GGTGCAGGAA	GAAATCCGGG	1200
AGGGGTTGTA	AATGTTGGAG	CTGATATCAA	GAAAACAATG	GAGGGGCCGG	TGGAGGGCAG	1260
AGACACAGCA	GAGCTTCCAG	CCCGCACATC	CCCCATGCCT	GGACACCCCCA	CTGCCAGCCC	1320
TACCTCCAGT	GAAGTCCAGC	AGGTGCCAAG	CCCTGTCTCC	TCTGAGCCTC	CCAAAGCTGC	1380
CAGACCCCT	GTGACACCTG	TCCTGCTAGA	GAAGAAAAGC	CCACTGGGCC	AGAGCCAGCC	1440
CACGGTGGCA	GGACAGCCCT	CAGCCCCCCC	AGCAGCAGAG	GAATATGGCT	ACATCGTCAC	1500
TGATCAGAAC	CCCCTGAGCC	TGGCTGCAGG	AGTGAAGCTG	CTGGAGATCC	TGGCTGAGCA	1560
TGTGCACATG	TCCTCAGGCA	GCTTCATCAA	CATCACTGTG	GTGGGACCAAG	CCCTCACCTT	1620
CCGCATCCGG	CACAATGAGC	AGAACCTGTC	TTTGGCTGAT	GTGACCCAAC	AAGCAGGGCT	1680
GGTGAAGTCT	GAACCTGAAAG	CACAGACAGG	GCTCCAAATC	TTGCAGACAG	GAGTGGGACA	1740
GAGGGAGGAG	GCAGCTGCAG	TCCTTCCCCA	AACTGCGCAC	AGCACCTCAC	CCATGCGCTC	1800
AGTGCCTGCTC	ACTCTGGTGG	CCCTGGCAGG	TGTGGCTGGG	CTGCTGGTGG	CTCTGGCTGT	1860
GGCTCTGTGT	GTGCGGCAGC	ATGCGCGCA	GCAAGACAAG	GAGCGCCTGG	CAGCCCTGGG	1920
GCCTGAGGGG	CCCCATGGTG	ACACTACCTT	TGAGTACCAAG	GACCTGTGCC	GCCAGCACAT	1980
GGCCACGAAG	TCCTTGTCA	ACCGGGCAGA	GGGTCCACCG	GAGCCTTCAC	GGGTGAGGAG	2040
TGTGTCTCC	CAGTTCAGCC	ACGCAGCCCA	GGCCAGCCCC	AGCTCCCACA	GCAGCACCCC	2100
GTCTGGTGC	GAGGAGCCGG	CCCAAGCCAA	CATGGACATC	TCCACGGGAC	ACATGATTCT	2160
GGCATAATCG	GAGGATCACC	TGCGGAACCG	GGACCCCTT	GCCAAGGAGT	GGCAGGCCCT	2220
CTGTGCCTAC	CAAGCAGAGC	AAACACCTG	TGCCACCGCG	CAGGGGGAGG	GCAACATCAA	2280

29

AAAAGAACCGG CATCCTGACT TCCTGCCCTA TGACCATGCC CGCATAAAAC TGAAGGTGGA	2340
GAGCAGCCCT TCTCGGAGCG ATTACATCAA CGCCAGCCCC ATTATTGAGC ATGACCCTCG	2400
GATGCCAGCC TACATAGCCA CGCAGGGCCC ECTGTCCCAT ACCATCGCAG ACTTCTGGCA	2460
GATGGTGTGG GAGAGCGGCT GCACCGTCAT CGTCATGCTG ACCCCGCTGG TGGAGGATGG	2520
TGTCAAGCAG TGTGACCGCT ACTGGCCAGA TGAGGGTGCC TCCCTCTACC ACGTATATGA	2580
GGTGAACCTG GTGTCGGAGC ACATCTGGTG CGAGGACTTT CTGGTGCGGA GCTTCTACCT	2640
GAAGAACGTG CAGACCCAGG AGACGCGCAC GCTCACCCAG TTCCACTTCC TCAGCTGCC	2700
GGCAGAGGGC ACACCGGCCT CCACCGGCC CCTGCTGGAC TTCCGCAGGA AGGTGAACAA	2760
GTGCTACCGG GGCGCCTCCT GCCCCATCAT CGTGCAC TGC AGTGATGGTG CGGGGAGGAC	2820
CGGCCACCTAC ATCCTCATCG ACATGGTCTT GAACCCATG GCAAAAGGAG TGAAGGAGAT	2880
TGACATCGCT GCCACCCCTGG AGCATGTCCG TGACCCAGGG CCTGGCCTTG TCCGCTCTAA	2940
GGACCAGTTT GAATTGCCC TGACAGCCGT GCGGGAGGAA GTGAATGCCA TCCTCAAGGC	3000
CCTGCCCTAG TGAGACCCCTG GGGCCCTTG GCGGGCAGCC CAGCCTCTGT CCCTCTTGC	3060
CTGTGTGAGC ATCTCTGTGT ACCCACTCCT CACTGCCCA CCAGCCACCT CTTGGCATG	3120
CTCAGCCCTT CCTAGAAAGAG TCAGGAAGGG AAAGCCAGAA GGGGCACGCC TGCCCAGCCT	3180
CCCATCCCAG AGCCTGGGC ATCCCAGAGC CCAGGGCATC CCATGGGGGT GCTGCAGCCA	3240
GGAGGAGAGG AAAGGACATG GGTAGCAATT CTACCCAGAG CCTTCTCCTG CCTACATTCC	3300
CTGGCCTGGC TCTCCTGTAG CTCTCCTGGG GTTCTGGAG TTCCCTGAAC ATCTGTGTGT	3360
GTCCCCCTAT GCTCCAGTAT GGAAGAATGG GCTGGAGGGT CGCCACACCC GGCTCCCCCT	3420
GCTTCTCAGC CCCGGGCCTG CCTCTGACTC ACACCTGGGC GCTCTGCCCT CCCTGGCCTC	3480
ACGCCAGCC TGGTCCCACC ACCCTCCCAC CATGCGCTGC TCAACCTCTC TCCTTCTGGC	3540
GCAAGAGAAC ATTTCTAGAA AAAACTACTT TTGTACCAAGT GTGAATAAAG TTAGTGTGTT	3600
GTCTGTGCAG CTG	3613

## (2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4992 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CTCGAGGGGC CTAGACATTG CCCTCCAGAG AGAGCACCCA ACACCCCTCCA GGCTTGACCG	60
GCCAGGGTGT CCCCTTCCTA CCTTGGAGAG AGCAGCCCCA GGGCATCCTG CAGGGGGTGC	120
TGGGACACCA GCTGGCCTTC AAGGTCTCTG CCTCCCTCCA GCCACCCAC TACACGCTGC	180

TGGGATCCTG GATCTCAGCT CCCTGGCCGA CAAACACTGGC AAAACTCCTAC TCATCCACGA	240
AGGCCCTCCT GGGCATGGTG GTCTTCCCAGC CTGTTCTCA CACACCTTGT	300
TAGTGCOCAG CCCCTGAGGT TGCAGCTGGG GGTGTCCTG AAGGGCTGTG AGCCCCCAGG	360
AAGCCCTGGG GAAGTGCCTG CCTTGCTCC CCCCCGGCCCT CCCAGGCCCT GGCTCTGCC	420
TCCTACCTGG GCTCCCCCA TCCAGCCTCC CTCCCTACAC ACTCCCTCTCA AGGAGGCACC	480
CATGTCTCT CCAGCTGCCG GGCCTCAGAG CACTGTGGCG TCCTGGGCA GCCACCGCAT	540
GTCCTGCTGT CGCATGGCTC AGGGTGGAAA GGGCGGAAGG GAGGGGTCT GCAGATAGCT	600
GGTCCCCACT ACCAAACCCG CTCGGGCAG GAGAGCCAAA GGCTGGGTGT GTGCAGAGCG	660
GCCCCGAGAG GTTCCGAGGC TGAGGCCAGG GTGGGACATA GGGATGCGAG GGGCCGGGGC	720
ACAGGATACT CCAACCTGCC TGCCCCATG GTCTCATCCT CCTGCTTCTG GGACCTCCTG	780
ATCCTGCCCT TGGTCTAAG AGGCAGGTAA GGGGCTGCAG GCAGCAGGGC TCGGAGCCA	840
TGCCCCCTCA CCATGGGTCA GGCTGGACCT CGAGGTGCCT GTTCTGGGA GCTGGGAGGG	900
CCGGAGGGGT GTACCCCAAGG GGCTCAGCCC AGATGACACT ATGGGGGTGA TGGTGTCTG	960
GGACCTGCCG AGGAGAGGGG AGATGGGCTC CCAGAAGAGG AGTGGGGGCT GAGAGGGTGC	1020
CTGGGGGCC AGGACGGAGC TGGGCCAGTG CACAGCTTCC CACACCTGCC CACCCCCAGA	1080
GTCCTGCCGC CACCCCCAGA TCACACGGAA GATGAGGTCC GAGTGGCCTG CTGAGGACTT	1140
GCTGCTTGTGCCC CAGGTCTGC CCTCCTTCTG CCACCCCTGGG GAGCTGAGGG	1200
CCTCAGCTGG GGCTGCTGTC CTAAGGCAGG GTGGGAACTA GGCAAGCCAGC AGGGAGGGGA	1260
CCCCCTCCCTC ACTCCCACCTC TCCCACCCCC ACCACCTTGG CCCATCCATG GCGGCATCTT	1320
GGGCCATCCG GGACTGGGGA CAGGGGTCTT GGGGACAGGG GTCCGGGGAC AGGGTCTGG	1380
GGACAGGGGT GTGGGGACAG GGGTCTGGGG ACAGGGGTGT GGGGACAGGG GTGTGGGAC	1440
AGGGTCTGG GGACAGGGGT GTGGGGACAG GGGTCTGGGG ACAGGGGTGT GGGGACAGGG	1500
GTCCTGGGAC AGGGGTGTGG GGACAGGGGT GTGGGGACAG GGGTCTGGGG ACAGGGGTGT	1560
GGGGACAGGG GTCCCTGGGA CAGGGGTGTG GGGACAGGGG TGTGGGGACA GGGGTGTGG	1620
GACAGGGGTG TGGGGACAGG GTCTCTGGGG ATAGGGGTGT GGGGACAGGG GTGTGGGAC	1680
AGGGGTCCCG GGGACAGGGG TGTGGGGACA GGGTGTGGG GACAGGGCTC CTGGGGACAG	1740
GGGTCTGAGG ACAGGGGTGT GGGCACAGGG GTCTCTGGGA CAGGGGTCTT GGGGACAGGG	1800
GTCCTGGGGA CAGGGGTCTG GGGACACCCAG CGCAAAGAGC CCCGCCCTGC AGGCTCCAGC	1860
TCTCTGGTC TAATCTGGAA AGTGGCCCAG GTGAGGGCTT TGCTCTCCTG GAGACATTG	1920
CCCCCAGCTG TGAGCAGGGA CAGGTCTGGC CACCGGGCCC CTGGTTAAGA CTCTAATGAC	1980
CCGCTGGTCC TGAGGAAGAG GTGCTGACGA CCAAGGAGAT CTTCCACAG ACCCAGCACC	2040
AGGGAAATGG TCCGGAAATT GCAGCCTCAG CCCCCAGCCA TCTGCCGACC CCCCCACCCC	2100
GCCCTAACATGG GCCAGGGCGGC AGGGGTTGAC AGTAGGGGA GATGGGCTCT GAGACTATAA	2160

AGCCAGCGGG	GGCCCAGCAG	CCCTCAGCCC	TCCAGGACAG	GCTGCATCAG	AAGAGGCCAT	2220
CAAGCAGGTC	TGTTCCAAGG	GCCTTGCCT	CAGGTGGCT	CAGGGTTCCA	GGGTGGCTGG	2280
ACCCCAGGCC	CCAGCTCTGC	AGCAGGGAGG	ACGTGGCTGG	GCTCGTGAAG	CATGTGGGG	2340
TGAGCCCAGG	GGCCCCAAGG	CAGGGCACCT	GGCCTTCAGC	CTGCCTCAGC	CCTGCCTGTC	2400
TCCCAGATCA	CTGTCCCTCT	GCATGGCCC	TGTGGATGCG	CCTCCCTGCC	CTGCTGGCGC	2460
TGCTGGCCCT	CTGGGGACCT	GACCCAGCCG	CAGCCTTTGT	GAACCAACAC	CTGTGCGGCT	2520
CACACCTGGT	GGAAGCTCTC	TACCTAGTGT	GCAGGGAAACG	AGGCTTCTTC	TACACACCCA	2580
AGACCCGCG	GGAGGCAGAG	GACCTGCAGG	GTGAGCCAAC	CGCCCATTCG	TGCCCCTGGC	2640
CGCCCCCAGC	CACCCCCCTGC	TCCTGGCGCT	CCCACCCAGC	ATGGGCAGAA	GGGGGCAGGA	2700
GGCTGCCACC	CAGCAGGGGG	TCAGGTGCAC	TTTTTTAAAA	AGAAGTTCTC	TTGGTCACGT	2760
CCTAAAAGTG	ACCAGCTCCC	TGTGGCCOCAG	TCAGAATCTC	AGCCTGAGGA	CGGTGTTGGC	2820
TTCGGCAGCC	CCGAGATAACA	TCAGAGGGTG	GGCACGCTCC	TCCCTCCACT	CGCCCOCTCAA	2880
ACAAATGCC	CGCAGCCCAT	TTCTCCACCC	TCATTGATG	ACCGCAGATT	CAAGTGT	2940
GTTAAGTAAA	GTCCTGGGTG	ACCTGGGGTC	ACAGGGTGCC	CCACGCTGCC	TGCCTCTGGG	3000
CGAACACCCC	ATCACGCCG	GAGGAGGGCG	TGGCTGCCTG	CCTGAGTGGG	CCAGACCCCT	3060
GTCGCCAGCC	TCACGGCAGC	TCCATAGTCA	GGAGATGGGG	AAGATGCTGG	GGACAGGCC	3120
TGGGGAGAAG	TACTGGGATC	ACCTGTTCA	GCTCCCACTG	TGACGCTGCC	CCGGGGGGGG	3180
GGAAGGAGGT	GGGACATGTG	GGCGTTGGGG	CCTGTAGGTC	CACACCCAGT	GTGGGTGACC	3240
CTCCCTCTAA	CCTGGGTCCA	GCCCCGGCTGG	AGATGGGTGG	GAGTGCACCC	TAGGGCTGGC	3300
GGGCAGGCCG	GCACGTGTC	TCCCTGACTG	TGTCCCTCCTG	TGTCCCTCTG	CCTCGCCCGCT	3360
GTTCCCGAAC	CTGCTCTGCC	CGGCACGTCC	TGGCAGTGGG	GCAGGTGGAG	CTGGGGGGGG	3420
GCCCTGGTGC	AGGCAGCCTG	CAGCCCTTGG	CCCTGGAGGG	GTCCCTGCCAG	AAGCGTGGCA	3480
TTGTGGAACA	ATGCTGTACC	AGCATCTGCT	CCCTCTACCA	GCTGGAGAAC	TACTGCAACT	3540
AGACGCAGCC	TGCAGCCAGC	CCCACACCCG	CCGCCTCCTG	CACCGAGAGA	GATGGAATAA	3600
AGCCCTTGAA	CCAGCCCTGC	TGTGCCGTCT	GTGTGTCTTG	GGGGCCCTGG	GCCAAGCCCC	3660
ACTTCCCGGC	ACTGTTGTGA	GCCCCCTCCA	GCTCTCTCCA	CGCTCTCTGG	GTGCCACAG	3720
GTGCCAACGC	CAGGCAGGCC	CAGCATGCAG	TGGCTCTCCC	CAAAGCGGCC	ATGCCCTGTTG	3780
GCTGCCTGCT	GCCCCCACCC	TGTGGCTCA	GGTCCAGTAT	GGGAGCTTCG	GGGGTCTCTG	3840
AGGGGCCAGG	GATGGTGGGG	CCACTGAGAA	GTGACTCTGT	CAGTAGCCGA	CCTGGAGTCC	3900
CCAGAGACCT	TGTTCAGGAA	AGGGAATGAG	AACTTCCAG	CAATTTCCC	CCCACCTAGC	3960
CCTCCCAGGT	TCTATTTTA	GAGTTATTTC	TGATGGAGTC	CCTGTGGAGG	GAGGAGGCTG	4020
GGCTGAGGGA	GGGGGTCCCTG	CAGGGCGGGG	GGCTGGGAAG	GTGGGGAGAG	GCTGCCGAGA	4080
GCCACCCGCT	ATCCCCAGCT	CTGGGCAGCC	CCGGGACAGT	CACACACCC	GGCCTCGCGG	4140

32

CCCAAGCTGG CAGCCGTCTG CAGCCACAGC TTATGCCAGC CCAGGTCCAG CCAGACACCT	4200
GAGGGACCCA CTGGTGCCTT GGAGGAAGCA GGAGAGGTCA GATGGCACCA TGAGCTGGGG	4260
CAGGTGCAGG GACCGTGGCA GCACCTGGCA GGGCCTCAGA ACCCATGCCT TGGGCACCCC	4320
GGCCATGAGG CCCTGAGGAT TGCAGCCAA GAGAAGCAGG GAACGCCAGG GCCACAGGGG	4380
CAGAGACCAG GCCAGGGTCC CTTGCGGGCC TTAGCCCACC CCCTCCCACT AAGCAGGGGC	4440
TGCTTGGCTA GGCTTCCTTT TGCTACAGAC CTGCTGCTCA CCCAGAGGCC CACGGGCCCT	4500
AGTGACAAGG TCGTTGTGGC TCCAGGTCTT TGGGGTCCT GACACAGAGC CTCTTCTGCA	4560
GCACCCCTGA GGACAGGGTG CTCCGCTGGG CACCCAGCCT AGTGGGCAGA CGAGAACCTA	4620
GGGGCTGCCT GGGCCTACTG TGGCCTGGGA GGTCAGCGGG TGACCCCTAGC TACCCCTGTGG	4680
CTGGGCCACT CTGCCTGCCA CCCAGGCCAA ACCAATCTGC ACCTTTCCTG AGAGCTCCAC	4740
CCAGGGCTGG GCTGGGGATG GCTGGGCCTG GGGCTGGCAT GGGCTGTGGC TGCAGACCAC	4800
TGCCAGCTTG GGCCTCGAGG CCAGGAGCTC ACCCTCCAGC TGCCCCGCCT CCAGAGTGGG	4860
GGCCAGGGCT GGGCAGGCAG GTGGACGGCC GGACACTGGC CCCGGAAGAG GAGGGAGGCAG	4920
GTGGCTGGGA TCGGCAGCAG CCGTCCATGG GAACACCCAG CCGGCCCCAC TCGCACGGGT	4980
AGAGACAGGC GC	4992

## (2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 10 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS:  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: C-terminal

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Xaa Xaa Gly Ser His His His His His His		
1	5	10

## (2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 27 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid  
 (A) DESCRIPTION: /desc = "DNA for bridge peptide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

AAGAAGAAGC GCCCGCGAAA GAAGAAG

27

## (2) INFORMATION FOR SEQ ID NO:13:

33

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "DNA for bridge peptide"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AAGAAGAAGC GATCGCGAAA GAAGAAG

27